Engineer Data

Gjorde pipelinen för geodata 50 gånger snabbare

Dokumenterade en förbättring på 50 gånger i en geografisk datapipeline i grundarens arbete hos Utiligize genom SQL- och datamodelländringar i PostgreSQL, MS SQL och Google Cloud BigQuery.

2022 – 2024

50x

snabbare datapipelines för geodata

Situation. Organisationen hade betydande förseningar i bearbetningen av stora mängder geografiska data, vilket påverkade effektiviteten i datadrivna beslutsprocesser och analytisk rapportering.

Uppgift. Målet var att optimera pipelinen för geografiska data för att förbättra prestandan och korta bearbetningstiden, så att dataanalyser kunde utföras effektivare och i realtid.

Handling. De befintliga strukturerna för SQL-programmering och datamodellering i PostgreSQL, MS SQL och Google Cloud BigQuery analyserades noggrant, vilket blottlade flaskhalsar i form av ineffektiva indexeringsstrategier, suboptimalt utformade SQL-frågor och avsaknad av lämpliga integritetsvillkor. För att komma till rätta med problemen:

- Datamodellerna gjordes om för att normalisera kritiska dataset, och partitioneringsstrategier infördes för att korta tiderna för datahämtning.
- Tillämpade avancerade indexeringstekniker, bl.a. B-tree- och GiST-index för PostgreSQL, filtrerade index i MS SQL och klustring i BigQuery.
- Optimerade komplexa SQL-frågor genom att refaktorera underfrågor, minska antalet join-operationer och använda materialiserade vyer där det var relevant.
- Etablerade integritetsvillkor som främmande nycklar och CHECK-villkor för att säkerställa datakonsistens utan att ge avkall på prestandan.

Resultat. Grundarens resultatbeskrivning anger en förbättring på 50 gånger för den geografiska datapipelinen i detta uppdrag. Detta är det dokumenterade projektresultatet, inte en garanti för andra datamängder eller system. Den publicerade beskrivningen anger inte mätningens datamängd, maskinvara eller jämförelsemetod.

Använd erfarenheten i ert projekt

Använd ”Gjorde pipelinen för geodata 50 gånger snabbare” som utgångspunkt för en liknande utmaning i er organisation. Tjänsterna bakom arbetet finns tillgängliga för ert projekt. Hittar databasfrågorna och de saknade indexen bakom ett system som har blivit tyngre för varje år, och åtgärdar dem som spelar roll.

Se våra tjänster Hör av er →

Se bakom resultatet

Läs om situationen, besluten och resultatet i exemplet. Vi kan diskutera vad som passar era förutsättningar och vad som kräver ett annat angreppssätt. Gjorde pipelinen för geodata 50 gånger snabbare — 50x snabbare datapipelines för geodata

Se våra tjänster Hör av er →

Behöver dina datapipelines gå snabbare?

Prata om ert projekt med ingenjören som ansvarar för leveransen. Vi avtalar omfattning, hur priset beräknas och överlämning skriftligt innan arbetet börjar.

Roll i uppdraget: Dataingenjör. Det bygger på Databaser, PostgreSQL, SQL, Datalager, Data engineering, Datapipelines (ETL/ELT), Datastyrning, Prestandaoptimering och GIS / geodata. Och nu när ni är här: vi är ganska bra på GIS- och geodatalösningar, Optimering av databasprestanda, Utveckling av datapipelines (ETL/ELT), Design av datalager och Dataförberedelse för AI.

Hör av er Se våra tjänster

En del av **en längre historik** — och den finns i sin helhet här.

Varje post är skriven på samma sätt: situationen, uppgiften, vad vi gjorde och vad som förändrades.

Se hela historiken

Definiera förbättringen ni behöver

Ta med ett representativt dataset, frågorna det ska besvara och dagens flaskhals. Vi kan avtala mått på kvalitet, genomströmning och återställning.

Se vad det kostar Hör av er →

Behöver dina datapipelines gå snabbare?

Fråga om besluten bakom ”Gjorde pipelinen för geodata 50 gånger snabbare”. Ni talar med personer som kan förklara arbetet och bedöma ert eget projekt.

Se våra tjänster

welcome@engineer.company +45 50 10 61 84 Köpenhamn, Danmark

Ta med dig den här sidan

Visitkort

Omsorgsfullt skapad med kärlek av D.

CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N

Upphovsrätt © 2025 – nu · Engineer ApS · Alla rättigheter förbehållna.

Utskrivet från https://data.engineer.company/sv/portfolio/made-a-geographical-data-pipeline-50x-faster/